

# Sistema de Reconocimiento de Rostros

Martín Correa  
Franco Chichizola  
Armando E. De Giusti  
Marcelo Naiouf

L.I.D.I., Facultad de Informática, Universidad Nacional de La Plata  
La Plata, Argentina, 1900  
{francoch,degiusti, mnaiouf}@lidi.info.unlp.edu.ar

## RESUMEN

En este trabajo se realiza un análisis de los sistemas de reconocimiento de rostros. Para ello se presentan los conceptos básicos sobre los cuales nos basaremos para la tarea de reconocer rostros.

Posteriormente, se estudia e implementa un modelo de sistema de reconocimiento de rostros, al cual se le realiza una serie de pruebas para medir su tasa de reconocimiento.

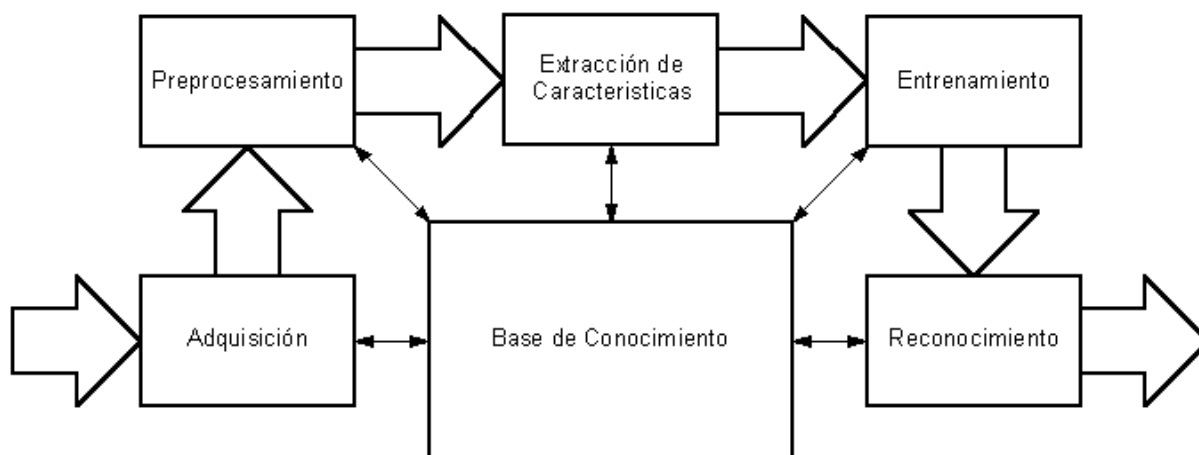
Es esperable que el sistema de reconocimiento de rostros a implementar, se comporte tolerante a imágenes de rostros con diferentes características de luminosidad, posiciones y expresiones.

La experimentación para verificar la metodología y el algoritmo propuesto utiliza una base de datos limitada de imágenes que han sido normalizadas internacionalmente.

Se obtienen excelentes resultados de reconocimiento “off line” y se indican las líneas de investigación tendientes al procesamiento paralelo del algoritmo, a fin de llegar a respuestas en tiempo real.

## INTRODUCCIÓN

El reconocimiento de rostros automatizado, involucra alguna de las siguientes etapas:



**Adquisición de la imagen:** En esta etapa, contamos con diversos medios de obtención de las imágenes digitalizadas de rostros. Los más utilizados son: videocámara, cámara digital o escaneado de fotografía entre otras. Se utilizaron para nuestro trabajo la base de datos ORL [9].

**Preprocesamiento de la imagen:** Dado que las imágenes de los rostros regularmente son tomadas en momentos diferentes, esto trae como consecuencia que las distintas imágenes de una misma persona tengan variaciones en cuanto a la iluminación, la orientación, y el tamaño del rostro. Por tal motivo, es necesario que la imagen sea preprocesada antes de que pueda ser utilizada. Entre las tareas de preprocesamiento más comunes podemos encontrar las siguientes: extraer la imagen del rostro de una imagen más grande que contenga información irrelevante para el reconocimiento; normalización en cuanto al tamaño, es decir, que todas las imágenes de los rostros tengan un tamaño similar, y la aplicación de algún método de filtrado para mejorar la calidad de la imagen. Las imágenes de la base de datos ORL, fueron transformadas en tamaño y formato gráfico.

**Extracción de Características:** La extracción de características es una de las etapas de la cual depende en gran medida el buen desempeño del sistema de reconocimiento de rostros. El objetivo principal de esta etapa es extraer la información más discriminante de un rostro, eliminando aquella que resulte irrelevante para el reconocimiento. En esta etapa podemos encontrar desde las técnicas que obtienen las características de manera manual, hasta técnicas mas sofisticadas que extraen las características de forma automática. En varias técnicas, las principales características que son extraídas de los rostros son aquellas que describen a cada uno de los componentes básicos del rostro tales como los ojos, la nariz, la boca, la barba, las cejas, y por supuesto, la relación que existe entre ellos, es decir, la distancia entre cada uno de estos componentes. La extracción de características la realizamos implícitamente tanto en el entrenamiento como en el reconocimiento.

**Entrenamiento:** El entrenamiento básicamente consiste en utilizar alguna forma de aprendizaje que le permita al sistema “aprender” los rostros que constituyen el conjunto de entrenamiento. El tipo de entrenamiento que sea utilizado para el aprendizaje, dependerá en gran medida, de la metodología que se esté utilizando para el reconocimiento.

**Reconocimiento:** Esta última etapa básicamente consiste en alimentar al sistema con imágenes de personas diferentes a las utilizadas durante el entrenamiento, esperando obtener como resultado, alguna forma de codificación que nos permita identificar de que persona se trata. (o la mayor similitud con las personas de la base de datos).

Para la implementación de este tipo de sistemas, independientemente de la técnica o metodología que se elija, se utilizan generalmente dos conjuntos de datos. El primer conjunto de datos es el que se emplea durante la etapa de aprendizaje, el cual es comúnmente llamado *conjunto de entrenamiento*. Se debe de tratar que los patrones que integran este conjunto, sean lo más diferente posible entre sí, y que además, representen al problema, para poder obtener un buen porcentaje de generalización. El segundo conjunto de patrones, es el que se utiliza durante la etapa de reconocimiento y es llamado *conjunto de prueba*.

## Resolución de algunos problemas complejos

El reconocimiento de rostros en condiciones “reales” resulta complejo, por varios factores que complican la aplicación directa de los algoritmos utilizados:

- ✓ localización del rostro en diferentes escenarios, que pueden ser muy disímiles.
- ✓ variaciones en la posición de los rostros y en la iluminación de la imagen.
- ✓ cambios en las expresiones de los rostros
- ✓ características especiales que pueden cambiar, como bigotes, barba, corte de pelo, anteojos, sombrero, etc.
- ✓ detección de rostro y seguimiento del mismo en condiciones de movimiento.
- ✓ extracción de características faciales y análisis de expresiones faciales
- ✓ categorización de rostros por raza, sexo o edad

## Tecnologías

Diversas tecnologías para el reconocimiento automatizado de rostros, han surgido en los últimos años. Entre las más destacadas encontramos el uso de Redes Neuronales [1][2] como ayuda al reconocimiento. También se observan técnicas recientes como el Análisis de Características Locales (Local Feature Analysis, LFA), el Modelo Markov (Hidden Markov Model, HMM) [3][4] o Elastic Bunch Graph Matching [5], en los cuales básicamente se hace énfasis en la localización de los elementos básicos de un rostro (ojos, nariz, boca, etc.), y su ubicación en el mismo.

Por otro lado, tenemos un modelo de reconocimiento basado en propiedades matemáticas de la imagen digitalizada, que capturan características invariantes de los rostros. Este es denominado modelo de Eigenfaces[6]. Básicamente consiste en utilizar imágenes de  $N \times N$ , como rostros de entrenamiento. Utilizando estos, se genera un rostro promedio, de tal forma que se podrá generar una matriz de diferencias. Posteriormente, se genera una matriz de covarianza, con la cual se realizara el cálculo del “Eigenspace” y de esta forma, lograr la generación de patrones.

El reconocimiento en este modelo, consistirá en encontrar el rostro cuyo patrón minimice la diferencia con el del rostro a reconocer.

## Análisis del sistema de reconocimiento a implementar

Es interesante estudiar y analizar la tecnología de reconocimiento de rostros mediante el uso de Eigenfaces, por los siguientes motivos:

- Otros modelos lo utilizan como piedra fundamental, derivando la técnica, o utilizándola como alimentación del modelo.
- Aparente simplicidad de implementación contra buenos resultados en grandes bases de datos
- Es una técnica resistente a variaciones de luminosidad, posición y características del rostro. No todas las técnicas tienen tolerancia a estas tres simultáneamente.
- Se realiza bajo un proceso netamente automático, en comparación a otros que necesitan (sobre todo en la etapa de entrenamiento) intervención humana

## **EIGENFACES**

Un gran número de imágenes de rostros se coleccionan en una base de datos. Luego, combinando todas las imágenes, se crea un conjunto de eigenfaces, - arreglos de dos dimensiones de los rostros de áreas claras y áreas oscuras -, y mirando a lo que es común a los grupos de individuos y donde más difieren. Los Eigenfaces trabajan como 'componentes de rostro'. Así como un color puede ser creado, mezclando los colores primarios, una imagen puede ser construida agregando conjuntamente, con diferentes intensidades de luz, un número de eigenfaces relativamente pequeño (cerca de 100 es suficiente para identificar la mayoría de las personas, pero solo para una posición y una condición de luminosidad relativa al rostro). Para identificar un rostro, el programa compara las características de su eigenface, que son codificadas en números llamado template, con aquellos en la base de datos, seleccionando los rostros cuyos templates matchean el objetivo más cerca.

### **Implementación**

#### ***Entrenamiento***

Para este algoritmo, se comienza asumiendo que existe un conjunto de rostros de entrenamiento de dimensión  $N \times N$  cada uno:

$$\Gamma_1, \Gamma_2, \Gamma_3, \dots, \Gamma_M$$

Como primer paso, estas imágenes de  $N \times N$  son reordenadas como un vector de una dimensión  $1 \times N^2$  usando un ordenamiento por fila (es decir, tal cual un orden lexicográfico).

Se obtiene el rostro promedio de la siguiente manera:

$$\Psi = \frac{1}{M} \sum_{n=1}^M \Gamma_n$$

y luego se le resta a cada una de las imágenes de entrenamiento, de forma que obtenemos un nuevo conjunto de vectores:

$$\Phi_i = \Gamma_i - \Psi.$$

Posteriormente, se deben calcular los autovectores a partir de la matriz de covarianza. Debido a la gran dimensión de esta matriz, se obtiene la matriz de covarianza reducida:

$$L = \frac{1}{M} A^T A$$

donde :

$$A = [ \Phi_1 \quad \Phi_2 \quad \dots \quad \Phi_M ].$$

que es un sistema de  $M \times M$ . Luego, se obtienen los autovectores de  $L$  (ordenados de mayor a menor de acuerdo a sus correspondientes autovalores), y se los pone en la matriz  $v$  (cada columna

es un autovector). Los eigenfaces  $\mathbf{u}$  (autovectores de la matriz de covarianza) se calculan como una combinación lineal de las columnas de  $\Phi_i$  y las columnas de  $\mathbf{v}$ , de la siguiente manera:

$$\mathbf{u}_\ell = \sum_{k=1}^M v_{\ell k} \Phi_k.$$

Luego se obtienen los patrones que representaran a cada rostro de la base, lo cual se hace de así:

$$\omega_k = \mathbf{u}_k^T \Phi_i \quad k = 1, 2, \dots, M'$$

formando el vector

$$\Omega^T = [\omega_1 \quad \omega_2 \quad \dots \quad \omega_{M'}]$$

que es el patrón correspondiente a la imagen  $i$ .

### ***Reconocimiento***

El reconocimiento de un rostro, luego, consiste en encontrar el rostro del conjunto de entrenamiento que minimiza la “distancia de rostro” con respecto a la imagen de entrada, es decir, que rostro de la base se asemeja al rostro de entrada al sistema. Descripto matemáticamente tendremos que:

$$\min \left( \|\Omega - \Omega_i\|^2 \right) \quad i = 1, 2, \dots, M.$$

La forma en que un rostro minimiza con otro, será variada, aunque podemos mencionar la distancia euclídea como la más elemental, y alguna otra técnica más sofisticada como mixture-distance.

Detectada la imagen que más se aproxima a la de entrada, de acuerdo a la distancia entre sus patrones se determina si se considera que es el mismo sujeto, o bien que no está en la base.

## **PRUEBAS REALIZADAS**

La base de datos de Rostros utilizada (llamada por sus creadores 'La Base de Datos de Rostros ORL' [9]), contiene un conjunto de 40 personas y 10 imágenes de rostro de cada una de ellas, tomadas entre Abril de 1992 y Abril de 1994 en el laboratorio AT&T de Cambridge.

Se realizaron una serie de pruebas para estudiar los resultados obtenidos por el sistema para distintas bases de conocimiento. Para la creación de las bases de conocimiento, se fueron realizando subconjuntos de una base de rostros que incluye únicamente una imagen de rostro por persona, es decir, 40 imágenes de rostro en total. Es así que se conformaron bases de conocimiento incluyendo cinco, diez, quince, veinte, veinticinco, treinta, treinta y cinco y cuarenta imágenes. Es evidente que no se han realizado para las pruebas, todas las bases de conocimiento posibles ya que con  $n$  imágenes, la cantidad posible de bases de conocimiento es:

$$\binom{n}{40} = \frac{n!}{40!(n-40)!}$$

En cuanto a las imágenes utilizadas para el reconocimiento formal, se tomaron los 400 rostros que posee la base de datos para testear contra la base de conocimiento.

Por otro lado se definió un “nivel de error” respecto al reconocimiento, cuyos valores pueden variar entre 0 y 10. Las bases de conocimiento fueron testeadas para cada uno de estos valores, para poder observar como se comporta el sistema en cada caso.

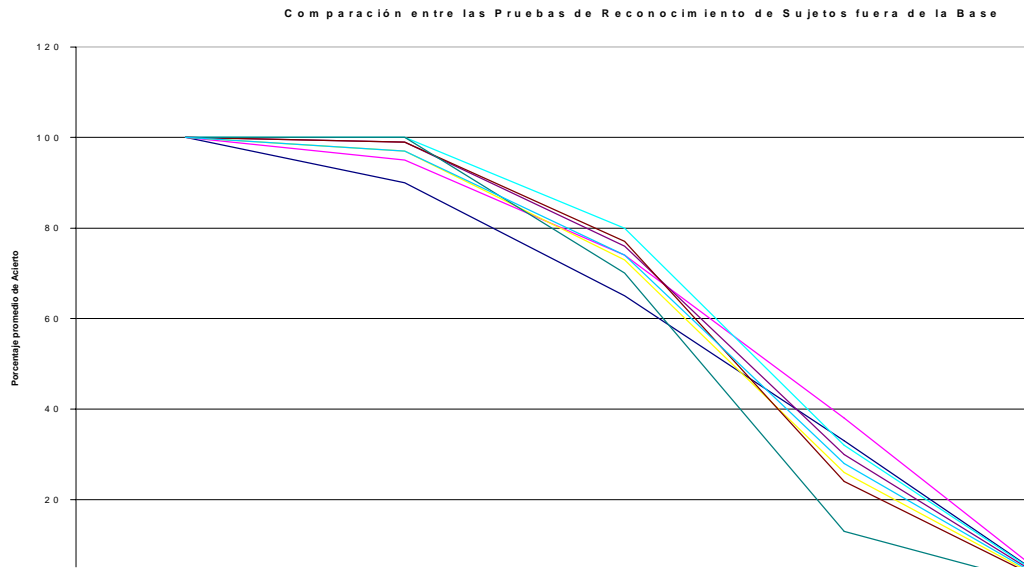
Cuando se utiliza un nivel de error 0, el rostro a reconocer debe ser casi idéntico al rostro que le corresponde (la misma persona con alguna variación de las anteriormente descritas) dentro de la base de conocimiento, para obtener un resultado positivo, es decir, lo reconozca correctamente; a medida que se aumenta el nivel de error, se permite una mayor diferencia entre el rostro a reconocer y el rostro que le corresponde. Cuando utilizamos un nivel de error 10, siempre obtendremos un resultado positivo (debido a que siempre devuelve el rostro que más se aproxime, asume un correcto reconocimiento).

Como consecuencia de esto, a medida que se aumenta el nivel de error, aumenta la cantidad de aciertos en el reconocimiento de los rostros de personas que poseen una imagen de rostro en la base de conocimiento, y disminuye la de las personas que no están en la base, hasta llegar a 0, con el nivel 10.

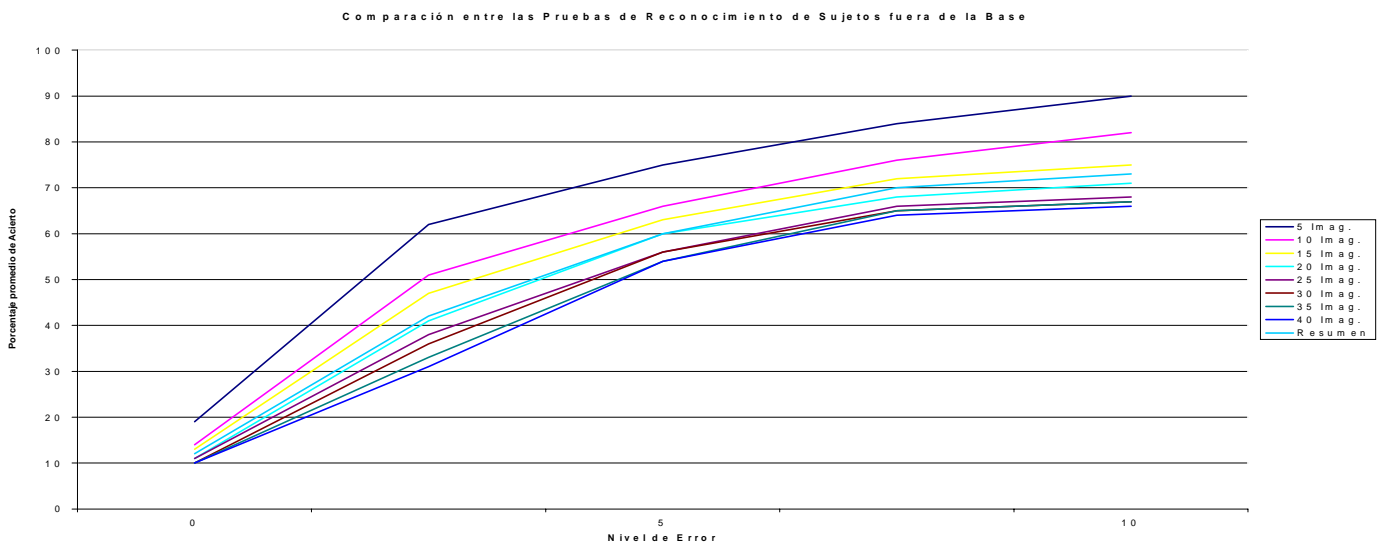
Cabe destacar que un “acierto” en el reconocimiento para una imagen de un rostro de una persona que posee imagen en la base de conocimiento, corresponde a indicar como resultado esa imagen de la base de conocimiento, mientras que para una imagen de rostro de una persona que no se encuentra en la base de conocimiento, un “acierto” corresponde a informar que no existe en la base. Es por ello que los resultados pueden observarse y/o analizarse mejor, ya que no existe una dependencia con la cantidad de imágenes de rostros en la base de conocimiento.

A continuación se provee un resumen de los resultados:

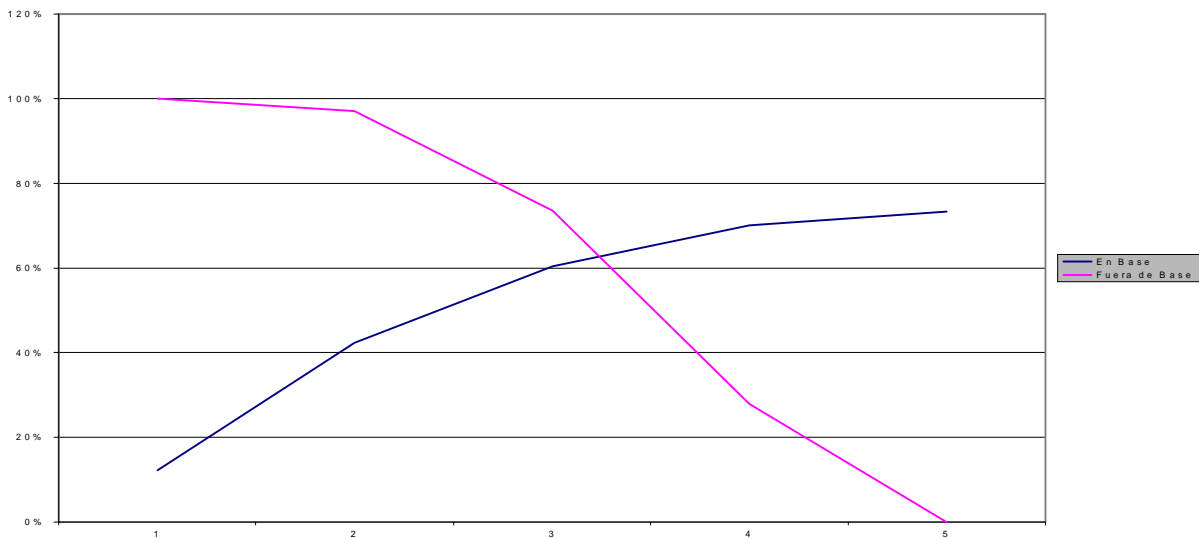
Tamaño de la base	Personas en la base					Personas fuera de la base				
	0	2	5	8	10	0	2	5	8	10
<b>5</b>	19%	62%	75%	84%	90%	100%	90%	65%	33%	0%
<b>10</b>	14%	51%	66%	76%	82%	100%	95%	74%	38%	0%
<b>15</b>	13%	47%	63%	72%	75%	100%	97%	73%	26%	0%
<b>20</b>	11%	41%	60%	68%	71%	100%	100%	80%	32%	0%
<b>25</b>	11%	38%	56%	66%	68%	100%	99%	76%	30%	0%
<b>30</b>	10%	36%	56%	65%	67%	100%	99%	77%	24%	0%
<b>35</b>	10%	33%	54%	65%	67%	100%	100%	70%	13%	0%
<b>40</b>	10%	31%	54%	64%	66%	-	-	-	-	-
<b>PROMEDIO T</b>	<b>12%</b>	<b>42%</b>	<b>60%</b>	<b>70%</b>	<b>73%</b>	<b>100%</b>	<b>97%</b>	<b>74%</b>	<b>28%</b>	<b>0%</b>



Notése la disminución del porcentaje de aciertos cuando se trata de sujetos fuera de la Base de conocimiento, a medida que se aumenta el nivel de error, tendiendo a cero, sin importar la cantidad de rostros en el conjunto de entrenamiento.



En este caso, para los rostros de prueba de personas que están en el conjunto de entrenamiento, el porcentaje de aciertos tiende a aumentar con el nivel de error. Observar como la curva se “suaviza” a medida que se aumenta el nivel de error



Tomando los valores totales de las pruebas de reconocimiento, se puede observar en este gráfico que un valor 5 para el nivel de error, nos conduce a tener una tasa media de reconocimiento de los rostros.

## **CONCLUSIONES Y LINEAS DE TRABAJO ACTUALES**

En líneas generales, se puede esperar que el sistema se comporte, tolerante a imágenes de rostros con diferentes características de luminosidad, posiciones y expresiones, tal cual se requería, logrando identificar los rostros de la base de conocimiento que se le enseñe, de manera acertada. Es interesante, también el hecho que se provee un nivel de precisión (que hemos denominado nivel de error), con el cual se puede realizar un análisis del comportamiento del sistema frente a las variaciones en el rostro, tantas veces mencionadas.

Se puede explicitar, luego de estas consideraciones que la respuesta en líneas generales de sistema es altamente eficiente, por lo que se consideran cumplidas las expectativas que abarcaron el desarrollo de esta aplicación.

En cuanto a futuros desarrollos que se pueden establecer a partir del estado actual de este trabajo, incluyendo tanto aspectos teóricos como de implementación.

- ✓ Utilizar el modelo Eigenface en un reconocimiento de rostros sobre vídeo.
- ✓ Mejorar el sistema para trabajar sobre imágenes de rostros en colores.
- ✓ Desarrollo de un sistema (o sub-sistema) normalizador de rostros potente, de tal forma que la base de datos de rostros pueda componerse de cualquier imagen que posea al menos un rostro.
- ✓ Implementación de otros modelos de sistemas de reconocimiento de rostros, haciendo hincapié en intentar automatizar etapas del mismo que actualmente se realizan manualmente.
- ✓ Paralelización del sistema para lograr respuestas en tiempo real [10].



## Bibliografía

- [1] S. Lin, S. kung, y L. Lin, "Face Recognition/Detection by Probabilistic Decision-Based Neural Network", *IEEE Trans. Neural Networks*, vol. 8, pp. 114-132, 1997.
- [2] Steve Lawrence, C. Lee Giles , Ah Chung Tsoi, Andrew D. Back "Face Recognition: A Convolutional Neural Network Approach". *IEEE Transactions on Neural Networks, Special Issue on Neural Networks and Pattern Recognition*, Volume 8, Number 1, pp. 98-113, 1997.
- [3] F. Samaria and S. Young, "HMM based architectures for face identification", *Image and Computer Vision*", vol. 12, pp 537-583, October 1994
- [4] F. Samaria, "Face Recognition Using Hidden Markov Models", PhD thesis, University of Cambridge, Cambridge, U.K. 1994.
- [5] Laurenz Wiskott, Jean-Marc Fellous, Norbert Krüger and Christoph von der Malsburg *Face Recognition by Elastic Bunch Graph Matching*, in *Intelligent Biometric Techniques in Fingerprint and Face Recognition*, eds. L.C. Jain et al., publ. CRC Press, ISBN 0-8493-2055-0, Chapter 11, pp. 355-396, (1999)
- [6] M. Turk and A. Pentland, "Face recognition using eigenfaces", in *Proceedings of International Conference on Pattern Recognition* , pp. 586-591,1991
- [7] Jose Gerardo Gonzalez, *Face Recognition and Detection Using Eigenfaces*,1997
- [8] MIT Media Laboratory Vision and Modeling Group (VISMODO) Face Recognition Demo Page  
<http://www-white.media.mit.edu/vismod/demos/facerec/>
- [9] AT&T Laboratories Cambridge, "The ORL Database of Faces",  
[ftp://ftp.uk.research.att.com/pub/data/att\\_faces.zip](ftp://ftp.uk.research.att.com/pub/data/att_faces.zip)  
<http://www.cam-orl.co.uk/facedatabase.html>
- [10] Champredonde Raúl y Chichizola Franco, "Paralelización de algoritmos en ADA sobre Clementina2. Aplicación a un sistema de reconocimiento de rostros". En Prensa.
- [11] Correa Martín Sebastián y Chichizola Franco. Trabajo de grado de Licenciatura en Informática "Sistema de Reconocimiento de Rostros". 2001.